

OIPE

## RAW SEQUENCE LISTING

DATE: 12/06/2001

PATENT APPLICATION: US/09/604,693A

TIME: 16:25:52

Input Set : D:\seqlistcorr.txt

Output Set: N:\CRF3\12062001\I604693A.raw

p.5

ENTERED

3 <110> APPLICANT: Pompejus, Markus  
 4 Kroger, Burkhard  
 5 Schroder, Hartwig  
 6 Zelder, Oskar  
 7 Haberhauer, Gregor

9 <120> TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS  
 10 INVOLVED IN DNA REPLICATION, PROTEIN SYNTHESIS,  
 11 AND PATHOGENESIS

13 <130> FILE REFERENCE: BGI-130CP  
 15 <140> CURRENT APPLICATION NUMBER: US 09/604,693A  
 16 <141> CURRENT FILING DATE: 2000-06-27  
 18 <160> NUMBER OF SEQ ID NOS: 432  
 20 <210> SEQ ID NO: 1  
 21 <211> LENGTH: 965  
 22 <212> TYPE: DNA  
 23 <213> ORGANISM: Corynebacterium glutamicum  
 25 <220> FEATURE:  
 26 <221> NAME/KEY: CDS  
 27 <222> LOCATION: (1)..(942)  
 28 <223> OTHER INFORMATION: RXN00625  
 30 <400> SEQUENCE: 1

31 ggc acc gat gaa gac tac atc cgc gtc gcc ttc atc tac gac caa acc 48  
 32 Gly Thr Asp Glu Asp Tyr Ile Arg Val Ala Phe Ile Tyr Asp Gln Thr  
 33 1 5 10 15  
 35 acc gtc aag ccc gtc ggc gaa tcc cga atc ttc gac gac gca gcc ttc 96  
 36 Thr Val Lys Pro Val Gly Glu Ser Arg Ile Phe Asp Asp Ala Ala Phe  
 37 20 25 30  
 39 acc ggc acc gca cgc cag cca ctc gca cag gaa ttc cag cca ctc aac 144  
 40 Thr Gly Thr Ala Arg Gln Pro Leu Ala Gln Glu Phe Gln Pro Leu Asn  
 41 35 40 45  
 43 gac agc gag aaa tcc ttc gtc ggc gta gtc aac cac ttc aag tcc aag 192  
 44 Asp Ser Glu Lys Ser Phe Val Gly Val Val Asn His Phe Lys Ser Lys  
 45 50 55 60  
 47 ggc tct gtc act cgt gga gac gcc gac acc ggc gac ggc caa ggc aac 240  
 48 Gly Ser Val Thr Arg Gly Asp Ala Asp Thr Gly Asp Gly Gln Gly Asn  
 49 65 70 75 80  
 51 aac gcc aac gtt cgc gtc gca cag gca cag gca ctc atc gac cac ctg 288  
 52 Asn Ala Asn Val Arg Val Ala Gln Ala Gln Ala Leu Ile Asp His Leu  
 53 85 90 95  
 55 gaa aac cag gac gac tgg gca tcc aag cca atc ttc atc ctc ggc gac 336  
 56 Glu Asn Gln Asp Asp Trp Ala Ser Lys Pro Ile Phe Ile Leu Gly Asp  
 57 100 105 110  
 59 acc aac tcc tac gcc aag gaa acc gcg atg acc acc ctt tac ggc gct 384  
 60 Thr Asn Ser Tyr Ala Lys Glu Thr Ala Met Thr Thr Leu Tyr Gly Ala  
 61 115 120 125  
 63 ggc tac acc aac atc gcc acc gaa ttc gac gct ggc tac agc tac cag 432  
 64 Gly Tyr Thr Asn Ile Ala Thr Glu Phe Asp Ala Gly Tyr Ser Tyr Gln

## RAW SEQUENCE LISTING

DATE: 12/06/2001

PATENT APPLICATION: US/09/604,693A

TIME: 16:25:52

Input Set : D:\seqlistcorr.txt

Output Set: N:\CRF3\12062001\I604693A.raw

```

65      130      135      140
67 ttc tcc ggc cgc att ggc agc ctc gac cac gca ctc ggc aac gaa gca 480
68 Phe Ser Gly Arg Ile Gly Ser Leu Asp His Ala Leu Gly Asn Glu Ala
69 145      150      155      160
71 gcc atg aag cac gtc atc gac gcc gag gtc tgg gac atc aac gct gac 528
72 Ala Met Lys His Val Ile Asp Ala Glu Val Trp Asp Ile Asn Ala Asp
73      165      170      175
75 gaa gca atc gca ttc gaa tac tcc cgt cga ctc aac aac acc tcc gac 576
76 Glu Ala Ile Ala Phe Glu Tyr Ser Arg Arg Leu Asn Asn Thr Ser Asp
77      180      185      190
79 gta ttc gag aac aac gtc ttc cgc tcc tcc gac cac gac ccg atc aag 624
80 Val Phe Glu Asn Asn Val Phe Arg Ser Ser Asp His Asp Pro Ile Lys
81      195      200      205
83 gtc gga ttc aac ctc agc gag acc act gag ccc acc att ccg gta gag 672
84 Val Gly Phe Asn Leu Ser Glu Thr Thr Glu Pro Thr Ile Pro Val Glu
85      210      215      220
87 ccc act gat cct gca gaa cct acc gat cca act acc cca gtt aag cca 720
88 Pro Thr Asp Pro Ala Glu Pro Thr Asp Pro Thr Thr Pro Val Lys Pro
89 225      230      235      240
91 act gat ccg gta gag acc acg gat cca tct gag cca acc gac cct gca 768
92 Thr Asp Pro Val Glu Thr Thr Asp Pro Ser Glu Pro Thr Asp Pro Ala
93      245      250      255
95 gaa cct act gat cca gct gaa cca act gac cct gag gaa acg aag aag 816
96 Glu Pro Thr Asp Pro Ala Glu Pro Thr Asp Pro Glu Glu Thr Lys Lys
97      260      265      270
99 cca gag gag ccg aag aac cct ggt tcc tcc aac gga agc tcc caa tac 864
100 Pro Glu Glu Pro Lys Asn Pro Gly Ser Ser Asn Gly Ser Ser Gln Tyr
101      275      280      285
103 gcc acc att gca gca atc atc gca gca atc cta ggt gcc att gct ttg 912
104 Ala Thr Ile Ala Ala Ile Ile Ala Ala Ile Leu Gly Ala Ile Ala Leu
105      290      295      300
107 gcc ttc cag ttc ttc cat tca agt tct aat taactcttag ggagtaatcc 962
108 Ala Phe Gln Phe Phe His Ser Ser Ser Asn
109 305      310
111 cct 965
114 <210> SEQ ID NO: 2
115 <211> LENGTH: 314
116 <212> TYPE: PRT
117 <213> ORGANISM: Corynebacterium glutamicum
119 <400> SEQUENCE: 2
120 Gly Thr Asp Glu Asp Tyr Ile Arg Val Ala Phe Ile Tyr Asp Gln Thr
121 1 5 10 15
123 Thr Val Lys Pro Val Gly Glu Ser Arg Ile Phe Asp Asp Ala Ala Phe
124 20 25 30
126 Thr Gly Thr Ala Arg Gln Pro Leu Ala Gln Glu Phe Gln Pro Leu Asn
127 35 40 45
129 Asp Ser Glu Lys Ser Phe Val Gly Val Val Asn His Phe Lys Ser Lys
130 50 55 60
132 Gly Ser Val Thr Arg Gly Asp Ala Asp Thr Gly Asp Gly Gln Gly Asn

```

## RAW SEQUENCE LISTING

DATE: 12/06/2001

PATENT APPLICATION: US/09/604,693A

TIME: 16:25:52

Input Set : D:\seqlistcorr.txt

Output Set: N:\CRF3\12062001\I604693A.raw

```

133 65          70          75          80
135 Asn Ala Asn Val Arg Val Ala Gln Ala Gln Ala Leu Ile Asp His Leu
136          85          90          95
138 Glu Asn Gln Asp Asp Trp Ala Ser Lys Pro Ile Phe Ile Leu Gly Asp
139          100          105          110
141 Thr Asn Ser Tyr Ala Lys Glu Thr Ala Met Thr Thr Leu Tyr Gly Ala
142          115          120          125
144 Gly Tyr Thr Asn Ile Ala Thr Glu Phe Asp Ala Gly Tyr Ser Tyr Gln
145          130          135          140
147 Phe Ser Gly Arg Ile Gly Ser Leu Asp His Ala Leu Gly Asn Glu Ala
148 145          150          155          160
150 Ala Met Lys His Val Ile Asp Ala Glu Val Trp Asp Ile Asn Ala Asp
151          165          170          175
153 Glu Ala Ile Ala Phe Glu Tyr Ser Arg Arg Leu Asn Asn Thr Ser Asp
154          180          185          190
156 Val Phe Glu Asn Asn Val Phe Arg Ser Ser Asp His Asp Pro Ile Lys
157          195          200          205
159 Val Gly Phe Asn Leu Ser Glu Thr Thr Glu Pro Thr Ile Pro Val Glu
160          210          215          220
162 Pro Thr Asp Pro Ala Glu Pro Thr Asp Pro Thr Thr Pro Val Lys Pro
163 225          230          235          240
165 Thr Asp Pro Val Glu Thr Thr Asp Pro Ser Glu Pro Thr Asp Pro Ala
166          245          250          255
168 Glu Pro Thr Asp Pro Ala Glu Pro Thr Asp Pro Glu Glu Thr Lys Lys
169          260          265          270
171 Pro Glu Glu Pro Lys Asn Pro Gly Ser Ser Asn Gly Ser Ser Gln Tyr
172          275          280          285
174 Ala Thr Ile Ala Ala Ile Ile Ala Ala Ile Leu Gly Ala Ile Ala Leu
175          290          295          300
177 Ala Phe Gln Phe Phe His Ser Ser Ser Asn
178 305          310
181 <210> SEQ ID NO: 3
182 <211> LENGTH: 362
183 <212> TYPE: DNA
184 <213> ORGANISM: Corynebacterium glutamicum
186 <220> FEATURE:
187 <221> NAME/KEY: CDS
188 <222> LOCATION: (1)..(339)
189 <223> OTHER INFORMATION: FRXA00625
191 <400> SEQUENCE: 3
192 tcc gac cac gac ccg atc aag gtc gga ttc aac ctc agc gag acc act 48
193 Ser Asp His Asp Pro Ile Lys Val Gly Phe Asn Leu Ser Glu Thr Thr
194 1 5 10 15
196 gag ccc acc att ccg gta gag ccc act gat cct gca gaa cct acc gat 96
197 Glu Pro Thr Ile Pro Val Glu Pro Thr Asp Pro Ala Glu Pro Thr Asp
198 20 25 30
200 cca act acc cca gtt aag cca act gat ccg gta gag acc acg gat cca 144
201 Pro Thr Thr Pro Val Lys Pro Thr Asp Pro Val Glu Thr Thr Asp Pro
202 35 40 45

```

## RAW SEQUENCE LISTING

DATE: 12/06/2001

PATENT APPLICATION: US/09/604,693A

TIME: 16:25:52

Input Set : D:\seqlistcorr.txt

Output Set: N:\CRF3\12062001\I604693A.raw

```

204 tct gag cca acc gac cct gca gaa cct act gat cca gct gaa cca act 192
205 Ser Glu Pro Thr Asp Pro Ala Glu Pro Thr Asp Pro Ala Glu Pro Thr
206      50      55      60
208 gac cct gag gaa acg aag aag cca gag gag ccg aag aac cct ggt tcc 240
209 Asp Pro Glu Glu Thr Lys Lys Pro Glu Glu Pro Lys Asn Pro Gly Ser
210 65      70      75      80
212 tcc aac gga agc tcc caa tac gcc acc att gca gca atc atc gca gca 288
213 Ser Asn Gly Ser Ser Gln Tyr Ala Thr Ile Ala Ala Ile Ile Ala Ala
214      85      90      95
216 atc cta ggt gcc att gct ttg gcc ttc cag ttc ttc cat tca agt tct 336
217 Ile Leu Gly Ala Ile Ala Leu Ala Phe Gln Phe Phe His Ser Ser Ser
218      100      105      110
220 aat taactcttag ggagtaatcc cct 362
221 Asn
225 <210> SEQ ID NO: 4
226 <211> LENGTH: 113
227 <212> TYPE: PRT
228 <213> ORGANISM: Corynebacterium glutamicum
230 <400> SEQUENCE: 4
231 Ser Asp His Asp Pro Ile Lys Val Gly Phe Asn Leu Ser Glu Thr Thr
232 1      5      10      15
234 Glu Pro Thr Ile Pro Val Glu Pro Thr Asp Pro Ala Glu Pro Thr Asp
235      20      25      30
237 Pro Thr Thr Pro Val Lys Pro Thr Asp Pro Val Glu Thr Thr Asp Pro
238      35      40      45
240 Ser Glu Pro Thr Asp Pro Ala Glu Pro Thr Asp Pro Ala Glu Pro Thr
241      50      55      60
243 Asp Pro Glu Glu Thr Lys Lys Pro Glu Glu Pro Lys Asn Pro Gly Ser
244 65      70      75      80
246 Ser Asn Gly Ser Ser Gln Tyr Ala Thr Ile Ala Ala Ile Ile Ala Ala
247      85      90      95
249 Ile Leu Gly Ala Ile Ala Leu Ala Phe Gln Phe Phe His Ser Ser Ser
250      100      105      110
252 Asn
256 <210> SEQ ID NO: 5
257 <211> LENGTH: 1768
258 <212> TYPE: DNA
259 <213> ORGANISM: Corynebacterium glutamicum
261 <220> FEATURE:
262 <221> NAME/KEY: CDS
263 <222> LOCATION: (101)..(1768)
264 <223> OTHER INFORMATION: RXN02943
266 <400> SEQUENCE: 5'
267 aactgccagg tcgaccaa tgcctgccaa gcagactccg aaaaacacgg gtaattcata 60
269 tggcttgtat ctaatccata ctgaacagag gacctctcct atg tct cgc att tct 115
270      Met Ser Arg Ile Ser
271      1      5
273 gcg cgc act ctg gca atc gca ctt gcc ggt gca acc gcg gcc agc ctg 163
274 Ala Arg Thr Leu Ala Ile Ala Leu Ala Gly Ala Thr Ala Ala Ser Leu

```

## RAW SEQUENCE LISTING

DATE: 12/06/2001

PATENT APPLICATION: US/09/604,693A

TIME: 16:25:52

Input Set : D:\seqlistcorr.txt

Output Set: N:\CRF3\12062001\I604693A.raw

275				10				15				20					
277	gca	gtt	gtt	cca	gca	gca	aca	gct	aat	cct	gcc	gga	acc	gct	cct	gtc	211
278	Ala	Val	Val	Pro	Ala	Ala	Thr	Ala	Asn	Pro	Ala	Gly	Thr	Ala	Pro	Val	
279				25				30				35					
281	atc	aac	gaa	atc	tac	gaa	ggc	ggt	gga	aac	agc	gga	tcg	ttg	ttc	tcc	259
282	Ile	Asn	Glu	Ile	Tyr	Glu	Gly	Gly	Gly	Asn	Ser	Gly	Ser	Leu	Phe	Ser	
283			40					45				50					
285	aac	gac	ttc	att	gag	ctc	tac	aac	cca	acc	tca	ggg	gac	att	tcc	ctc	307
286	Asn	Asp	Phe	Ile	Glu	Leu	Tyr	Asn	Pro	Thr	Ser	Gly	Asp	Ile	Ser	Leu	
287		55					60					65					
289	gac	ggt	tgg	agc	gtt	acc	tac	tac	gca	gcc	aac	ggt	aac	tcc	ggc	gga	355
290	Asp	Gly	Trp	Ser	Val	Thr	Tyr	Tyr	Ala	Ala	Asn	Gly	Asn	Ser	Gly	Gly	
291	70					75				80					85		
293	acc	aca	aac	ctg	acc	gga	aac	atc	cct	gcc	aac	ggt	tac	tac	ctc	atc	403
294	Thr	Thr	Asn	Leu	Thr	Gly	Asn	Ile	Pro	Ala	Asn	Gly	Tyr	Tyr	Leu	Ile	
295				90					95						100		
297	cag	caa	cgc	gca	ggc	agc	aac	aac	acc	ggc	gct	ctg	cct	acc	cca	gac	451
298	Gln	Gln	Arg	Ala	Gly	Ser	Asn	Asn	Thr	Gly	Ala	Leu	Pro	Thr	Pro	Asp	
299				105					110					115			
301	gcc	acc	ggt	aac	ttg	gca	atg	ggt	gcc	tcc	caa	gga	tca	gtt	gca	ctg	499
302	Ala	Thr	Gly	Asn	Leu	Ala	Met	Gly	Ala	Ser	Gln	Gly	Ser	Val	Ala	Leu	
303			120					125					130				
305	acc	gac	aac	tct	ggc	cta	acc	gct	gac	ctt	gtc	gga	ttc	ggt	ggc	acg	547
306	Thr	Asp	Asn	Ser	Gly	Leu	Thr	Ala	Asp	Leu	Val	Gly	Phe	Gly	Gly	Thr	
307		135					140					145					
309	tcc	atg	ttt	gaa	gga	aca	gct	gct	gca	cct	gag	acc	agc	aac	aaa	ttg	595
310	Ser	Met	Phe	Glu	Gly	Thr	Ala	Ala	Ala	Pro	Glu	Thr	Ser	Asn	Lys	Leu	
311	150					155				160					165		
313	tct	gtt	caa	cgc	aaa	gaa	gtt	ggc	gct	gac	tct	gat	aac	aac	tcc	gta	643
314	Ser	Val	Gln	Arg	Lys	Glu	Val	Gly	Ala	Asp	Ser	Asp	Asn	Asn	Ser	Val	
315				170					175						180		
317	gac	ttc	gag	act	gga	gct	cca	act	cca	acg	tcc	tcg	gga	gga	tcc	gct	691
318	Asp	Phe	Glu	Thr	Gly	Ala	Pro	Thr	Pro	Thr	Ser	Ser	Gly	Gly	Ser	Ala	
319				185					190						195		
321	cct	gtt	gac	cca	ggc	gag	cca	gaa	act	cca	gta	aac	cct	ggg	gaa	aca	739
322	Pro	Val	Asp	Pro	Gly	Glu	Pro	Glu	Thr	Pro	Val	Asn	Pro	Gly	Glu	Thr	
323			200					205					210				
325	gtc	tcc	atc	gca	caa	atc	caa	gga	acc	ggt	ctc	gct	acc	cca	ctc	gag	787
326	Val	Ser	Ile	Ala	Gln	Ile	Gln	Gly	Thr	Gly	Leu	Ala	Thr	Pro	Leu	Glu	
327		215				220						225					
329	ggt	cag	acc	gtc	acc	acc	gaa	ggt	att	gtc	act	gcc	gtt	tac	gca	gaa	835
330	Gly	Gln	Thr	Val	Thr	Thr	Glu	Gly	Ile	Val	Thr	Ala	Val	Tyr	Ala	Glu	
331	230					235				240					245		
333	ggt	ggc	ttc	aac	ggt	tac	tac	atc	cag	aca	cct	gga	tct	ggt	act	gca	883
334	Gly	Gly	Phe	Asn	Gly	Tyr	Tyr	Ile	Gln	Thr	Pro	Gly	Ser	Gly	Thr	Ala	
335				250					255						260		
337	cca	aag	gtt	gct	ggc	gac	gca	tcc	gac	ggc	atc	ttc	gtc	tac	gtg	gga	931
338	Pro	Lys	Val	Ala	Gly	Asp	Ala	Ser	Asp	Gly	Ile	Phe	Val	Tyr	Val	Gly	
339				265					270						275		

Use of n and/or Xaa has been detected in the Sequence Listing.  
 Review the Sequence Listing to insure a corresponding  
 explanation is presented in the <220> to <223> fields of  
 each sequence using n or Xaa.

## VERIFICATION SUMMARY

DATE: 12/06/2001

PATENT APPLICATION: US/09/604,693A

TIME: 16:25:53

Input Set : D:\seqlistcorr.txt

Output Set: N:\CRF3\12062001\I604693A.raw

L:17773 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:139  
L:17775 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:139  
L:17779 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:139  
L:17780 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:139  
L:17783 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:139  
L:17784 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:139  
L:17787 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:139  
L:17818 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:140  
L:17821 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:140  
L:17824 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:140  
L:28853 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:255  
L:28854 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:255  
L:28857 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:255  
L:28858 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:255  
L:28928 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:256  
L:28931 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:256